D.I. URAL

ATILIM UNIVERSITY 2024

A COMPARATIVE STUDY OF AUGMENTATION TECHNIQUES FOR DEFECT DETECTION IN PCB MANUFACTURING: IDENTIFYING THE BEST APPROACHES FOR LIMITED DATASETS

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

DOĞAN IRMAK URAL

A MASTER OF SCIENCE THESIS
IN
THE DEPARTMENT OF COMPUTER ENGINEERING

DECEMBER 2024

A COMPARATIVE STUDY OF AUGMENTATION TECHNIQUES FOR DEFECT DETECTION IN PCB MANUFACTURING: IDENTIFYING THE BEST APPROACHES FOR LIMITED DATASETS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

BY

DOĞAN IRMAK URAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF COMPUTER ENGINEERING

DECEMBER 2024

Approval of the Graduate School of Natural and Applied Sciences, Atılım University.

____

Prof. Dr. Ender KESKİNKILIÇ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of **Master of Science in Computer Engineering, Atılım University**.

____

Prof. Dr. Gökhan ŞENGÜL
Head of Department

This is to certify that we have read the thesis A COMPARATIVE STUDY OF AUGMENTATION TECHNIQUES FOR DEFECT DETECTION IN PCB MANUFACTURING: IDENTIFYING THE BEST APPROACHES FOR LIMITED DATASETS submitted by DOĞAN IRMAK URAL and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____

Asst. Prof. Dr. Arda SEZEN
Supervisor

Examining Committee Members:

Prof. Dr. Gökhan ŞENGÜL
Department of Computer Engineering,
Atılım University
____

Asst. Prof. Dr. Arda SEZEN
Department of Computer Engineering,
Atılım University
____

Asst. Prof. Dr. Ekrem Çağlar YILMAZ
Department of Computer Engineering,
Ankara Bilim University
____

Date: *December 27, 2024*

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Doğan Irmak, URAL

Signature:

ABSTRACT

A COMPARATIVE STUDY OF AUGMENTATION TECHNIQUES FOR DEFECT DETECTION IN PCB MANUFACTURING: IDENTIFYING THE BEST APPROACHES FOR LIMITED DATASETS

URAL, Doğan Irmak
M.S., Department of Computer Engineering
Supervisor: Dr. Arda SEZEN

December 2024, 59 pages

In the rapidly expanding field of electronics manufacturing, the detection of defects in printed circuit boards (PCBs) has emerged as a crucial aspect of quality assurance. Although convolutional neural networks (CNNs) and advanced deep learning models like YOLOv8 have significantly improved automated defect detection, their effectiveness is often constrained by the limited availability of large, high-quality datasets. This study examines the influence of various data augmentation techniques on enhancing the accuracy of defect detection in PCBs, with a particular focus on small object detection in scenarios where datasets are restricted in size. Through a series of controlled experiments, we systematically compared image-based and bounding box-based augmentation strategies. The results reveal that moderate augmentations, such as the addition of noise, slight rotations, and subtle color adjustments, considerably enhance detection performance. In contrast, more aggressive augmentations, including large rotations and flips, were found to negatively impact accuracy. These findings indicate that adopting well-balanced augmentation techniques can help overcome the limitations posed by small datasets, providing PCB manufacturers with a viable approach to improving defect detection accuracy without relying on extensive data. Ultimately, the study demonstrates that optimizing augmentation techniques leads to a significant improvement of a maximum 11% across key performance metrics, including precision, recall, mAP50, and F1 score, elevating detection accuracy from 88% to an exceptional 99%.

Keywords: Artificial Intelligence, Defect Detection, Dataset Augmentation, Printed Circuit Board, Image Processing

ÖZ

BASKI DEVRE KARTI (PCB) ÜRETİMİNDE HATA TESPİTİ İÇİN ARTTIRMA TEKNİKLERİNİN KARŞILAŞTIRMALI BİR ÇALIŞMASI: SINIRLI VERİ SETLERİ İÇİN EN İYİ YAKLAŞIMLARIN BELİRLENMESİ

URAL, Doğan Irmak

Bilgisayar Mühendisliği Bölümü

Danışman: Dr. Arda SEZEN

Aralık 2024, 59 Sayfa

Elektronik üretim sektörünün hızla genişlemesiyle birlikte, baskılı devre kartlarında (PCB'ler) hata tespiti, kalite güvencesinin kritik bir unsuru haline gelmiştir. Evrişimli sinir ağları (CNN'ler) ve YOLOv8 gibi derin öğrenme modelleri, otomatik hata tespitinde önemli ilerlemeler kaydetmiş olsa da, bu modellerin başarısı genellikle büyük ve yüksek kaliteli veri kümelerinin sınırlı olmasıyla engellenmektedir. Bu çalışma, veri kümelerinin sınırlı olduğu durumlarda, özellikle küçük nesne tespitine odaklanarak, farklı veri artırma tekniklerinin PCB'lerde hata tespit doğruluğunu nasıl etkilediğini incelemektedir. Bir dizi kontrollü deney yoluyla, görüntü tabanlı ve sınırlayıcı kutu (bounding box) tabanlı veri artırma stratejilerini sistematik olarak karşılaştırdık. Sonuçlar, orta düzeyde gürültü ekleme, hafif döndürmeler ve ince renk ayarları gibi dengeli artırma yöntemlerinin tespit performansını önemli ölçüde artırdığını göstermektedir. Buna karşılık, büyük döndürmeler ve ters çevirme gibi daha agresif artırma tekniklerinin doğruluğu olumsuz etkilediği bulunmuştur. Bu bulgular, dengeli artırma tekniklerinin küçük veri kümelerinin getirdiği sınırlamaları aşmaya yardımcı olabileceğini ve PCB üreticilerine, geniş veri setlerine ihtiyaç duymadan, hata tespit doğruluğunu artırmak için etkili bir yol sunduğunu göstermektedir. Sonuç olarak, çalışma, iyileştirilmiş artırma tekniklerinin, kesinlik,

duyarlılık, mAP50 ve F1 skoru gibi temel performans metriklerinde maksimum %11'lik önemli bir artış sağladığını ve tespit doğruluğunu %88'den olağanüstü bir şekilde %99'a çıkardığını göstermektedir.
Anahtar Kelimeler: Yapay Zeka, Kusur Tespiti, Veri Kümesi Arttırımı, Baskı Devre Kartı, Görüntü İşleme

To My Family…

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my mentor and supervisor, Dr. Arda SEZEN, for his exceptional guidance, steadfast support, immense patience, and continual encouragement throughout this journey.

Finally, I would also like to thank my family and my fiancée for their support over the last 3 years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PCB	Printed Circuit Board
YOLO	You Only Look Once
CNN	Convolutional Neural Network
CW	Clockwise
CCW	Counterclockwise
UD	Upside Down
AI	Artificial Intelligence
ResNet	Residual Network
IoU	Intersection over Union

CHAPTER 1

INTRODUCTION

With the rapid advancements in electronics manufacturing, ensuring the quality and reliability of printed circuit boards (PCBs) has become more crucial than ever [1]. The increasing complexity of electronic devices, alongside the rise in production volumes, has intensified the demand for efficient and precise defect detection methods. High-quality PCBs are fundamental to product reliability and play a key role in minimizing production costs while maintaining a competitive market position. Traditionally, manual inspection has been central to quality control in PCB manufacturing; however, these methods are increasingly inadequate in meeting the requirements of modern, high-speed production lines [2]. In response, automated quality control systems have evolved from basic image processing techniques to sophisticated Artificial Intelligence (AI) applications, particularly those employing convolutional neural networks (CNNs) and deep learning models [3]. Although AI-driven methods have markedly enhanced defect detection speed and accuracy, their success is heavily dependent on the availability of large, high-quality datasets.

The primary obstacle to advancing PCB defect detection is not merely model refinement but the scarcity of expansive, diverse datasets essential for effective model training. Despite enhanced manufacturing processes that have reduced defect occurrences, the time-intensive nature of manual inspections—often requiring up to 15 minutes per PCB—reveals inefficiencies, especially within large-scale production environments. As the PCB industry continues its accelerated growth, the demand for robust, automated defect detection solutions becomes ever more pressing.

This study redirects focus from merely improving model performance to addressing the core issue of dataset limitations. We propose that optimizing data augmentation techniques can significantly improve the detection of small component defects on PCBs, a challenge compounded by the limited availability of large datasets. In contrast to previous studies that primarily concentrate on identifying the most effective detection models, our research emphasizes the critical role of augmentation in producing reliable and representative datasets. Through a systematic comparison of various image-based and bounding box-based augmentation techniques, this study aims to establish a methodological framework for creating high-quality training data, which can lead to more accurate and efficient defect detection.

The implications of this research are particularly significant for large-scale PCB manufacturers, where a transition to automated quality control could substantially reduce inspection times and operational costs. As the industry scales, integrating advanced AI-driven defect detection systems—supported by optimized data augmentation methods—has the potential to render manual inspections obsolete, enabling real-time identification and segregation of defective PCBs. Our findings suggest that by implementing appropriate augmentation techniques, the challenges associated with limited datasets can be alleviated, facilitating the development of more robust and dependable defect detection systems.

This article is organized as follows: the Methodology section details the augmentation techniques and experimental setup utilized in our study; the Results and Discussion section analyzes our findings and their implications; and finally, the Conclusion provides a summary of the key insights and recommendations for future research directions.

1.1 Literature Review

The study conducted by Ural and Sezen is a systematic review [1] of AI-driven methods for defect detection in Printed Circuit Boards, addressing the industry's pressing need for speed and accuracy due to increasing production demands and design complexity. The findings reveal that AI techniques such as YOLO (You Only Look Once) and ResNet are highly favored for their capability to balance speed and precision, essential for high-throughput manufacturing environments. YOLO demonstrated both high accuracy and rapid processing, making it a popular choice for real-time defect detection systems. Other methods, like Feature Pyramid Networks (FPN) and Deep Feature Learning, showcased their potential for specific tasks, highlighting the diversity of approaches in the field.

A key insight from the review is the critical influence of dataset quality and size on the performance of AI models. Approximately 57% of studies used self-procured datasets, reflecting a trend toward tailoring datasets to specific defect detection needs. However, 33% of the reviewed studies employed data augmentation techniques to overcome the limitations of small datasets, which significantly improved model generalization and accuracy. This underscores the importance of augmentation in scenarios where data collection is constrained, making AI accessible for niche PCB designs and low-volume production environments.

The study also identified copper defects (e.g., short circuits, open circuits) and solder joint issues as the most frequently targeted defect types, comprising 70% of the total focus. This prioritization aligns with their significant impact on production reliability and functionality. Furthermore, the review noted an increasing interest in integrated defect detection systems capable of not only identifying but also classifying defects, which adds value in automating quality control and enabling preemptive production adjustments.

Overall, the findings indicate a paradigm shift in PCB manufacturing, where AI and machine learning are driving the transition from traditional inspection methods to adaptive, efficient, and scalable solutions. These advancements promise to enhance production efficiency, reduce error rates, and support the industry's growing need for precision and speed in an increasingly complex manufacturing landscape.

CHAPTER 2

METHODOLOGY

2.1 Purpose

This research aims to determine the most effective data augmentation techniques to enhance defect detection in printed circuit boards (PCBs) utilizing the YOLOv8 model, particularly under conditions of limited datasets [4]. YOLO's architectural characteristics—such as challenges with detecting small and overlapping objects, limitations in identifying multiple objects in complex environments, a trade-off between accuracy and real-time processing speed, and sensitivity to image resolution and intricate backgrounds—prompted us to conduct a structured investigation focused on data optimization and augmentation. This research approach addresses the shortage of comprehensive datasets in the field, aiming to fill a significant gap in the existing body of work.

Defect detection is a crucial component of quality control within PCB manufacturing. However, confidentiality concerns often deter companies from sharing detailed datasets on PCB defects, complicating the development of robust detection models. Therefore, identifying optimal augmentation techniques becomes essential to maximize the effectiveness of limited datasets and improve model accuracy.

The findings from this study are anticipated to provide substantial benefits to PCB manufacturers, quality control engineers, and the broader electronics industry by enabling faster, more precise defect detection. Additionally, these results will offer valuable insights for the academic community, particularly those focused on detecting defects in small-scale objects.

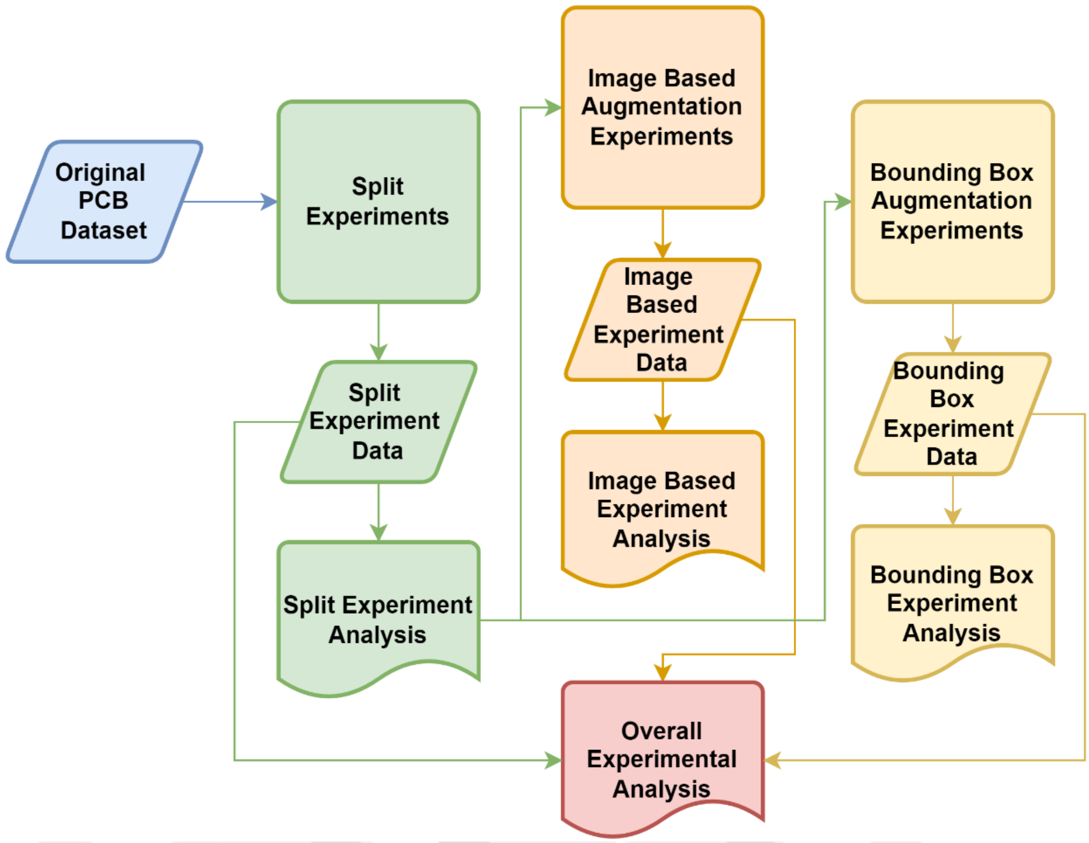

Figure 1 Overall Experimental Flow

2.2 Background Information About YOLO

YOLO, short for You Only Look Once, is an advanced computer vision technique used for object detection. Unlike image classification, where the goal is to recognize what is in an image, object detection also determines where the objects are within the image by drawing bounding boxes around them.

It is known and widely used because it is a real-time object detection framework that formulates detection as a single regression problem, predicting bounding box coordinates and class probabilities directly from an input image in a single forward pass through the network. By dividing the image into an M×M grid and assigning responsibility for detections to specific cells based on object center locations, YOLO avoids the computational overhead of region proposal-based methods.

It was selected for defect detection on small objects due to its combination of speed, accuracy, and ease of implementation. In the context of this thesis, the primary focus is not on the model itself but on comparing the effectiveness of various augmentation techniques. YOLO's straightforward setup and reliable performance made it the ideal choice to support this objective.

2.3 Research Questions and Hypotheses

Research Questions:

1. How do image-based augmentation techniques compare to bounding box-based techniques?
2. How does the amount of training data influence augmentation effectiveness?
3. Are certain augmentations more effective for detecting small components?
4. What is the impact of color augmentation techniques?

2.4 Research Design

This research adopts an experimental approach, systematically assessing and comparing various data augmentation techniques for defect detection in PCBs using YOLOv8. The experiments are conducted under controlled conditions, maintaining consistency by using the same YOLOv8 model and weight parameters throughout, ensuring that results are directly comparable across techniques.

Data Collection

Following the application of each augmentation technique, data is collected, and corresponding performance metrics—including precision, recall, mAP50, and F1 score—are documented. To prevent bias, randomization is integrated into the augmentation settings.

Statistical Analysis

The analysis centers on comparing precision, recall, mAP50, and F1 score across the various augmentation techniques.

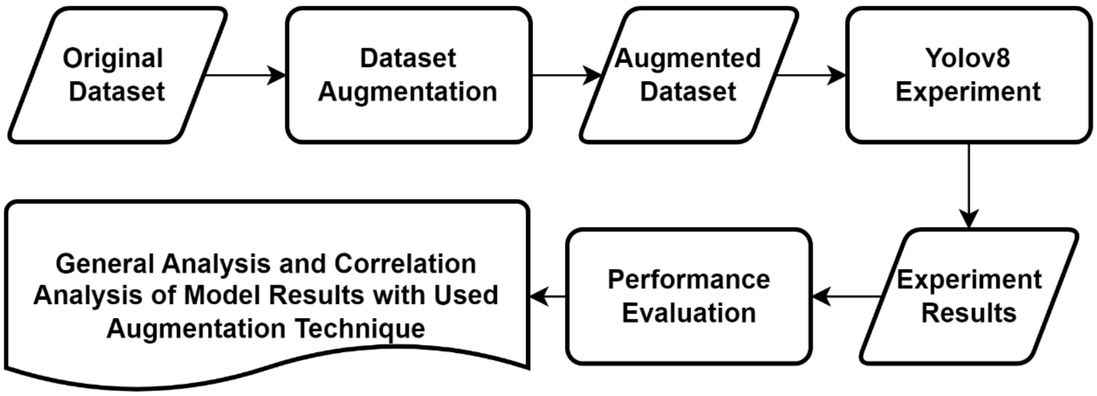

Figure 2 General Flow of the Research

2.5 Metrics Used in YOLO

In evaluating YOLO's performance, the following key metrics are employed to assess its accuracy and reliability in object detection tasks:

1. **Precision:**

Precision quantifies how many of the predicted bounding boxes are correct. It is calculated as:

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

Figure 3 Equation of Precision Metric

High precision indicates fewer false positives, ensuring predictions are accurate.

2. **Recall:**

Recall measures how many of the actual objects were correctly detected. It is expressed as:

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Figure 4 Equation of Recall Metric

High recall ensures the model identifies most of the actual objects.

3. mAP@50:

This metric evaluates the model's precision and recall at various confidence thresholds, focusing on Intersection over Union (IoU) ≥ 0.5 for a positive match. It calculates the area under the precision-recall curve, providing a comprehensive measure of detection performance.

4. F1 Score:

The F1 score is the harmonic mean of precision and recall, offering a balanced metric. It is computed as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 5 Equation of F1 Score Metric

A high F1 score indicates the model achieves a good balance between precision and recall.

These metrics collectively ensure a holistic evaluation of YOLO's object detection capabilities, highlighting its strengths and areas for improvement.

2.6 Materials and Equipment

2.6.1 Software

- **Google Colab Pro**: Employed to conduct all experiments on a virtual machine featuring an A100 GPU and 42GB of RAM.
- **YOLOv8 (v8.2.73)**: Chosen for its advanced performance in object detection tasks [5, 13], integrated through the ultralytics library.
- **Python 3.12.1**: Programming language used for executing the experiments.
- **Roboflow**: Applied for data augmentation, generating over 50,000 images from the original dataset.
- **Excel/Google Sheets**: Used for organizing and analyzing data.

2.6.2 Hardware

The experiments were performed on a Google Colab A100 virtual machine, offering the computational power required for deep learning tasks.

2.6.3 Dataset

The original dataset, titled "A PCB Dataset for Defects Detection and Classification" developed by Weibo Huang and Peng Wei at The Open Lab on Human-Robot Interaction, Peking University [5], was augmented using a variety of techniques. This process substantially expanded the dataset, providing a robust foundation for experimentation.

This dataset encompasses six primary copper defects commonly encountered on PCBs during the manufacturing process. These defects are detailed below:

- **Missing Hole**: A hole in the PCB that was not drilled or plated as intended, preventing proper component placement or connections.
- **Mouse Bit**: A small circular defect resembling a "bite," often caused by incomplete drilling or fabrication errors.
- **Open Circuit**: A break in the conductive path that disrupts the electrical flow, causing components to lose connection.
- **Short**: An unintended connection between two conductive paths, leading to electrical malfunction or damage.
- **Spur**: An unwanted, small, protruding conductor extending from a trace, which can cause shorts or signal issues.
- **Spurious Copper**: Unintended copper residues left on the PCB, which may cause shorts or interfere with functionality.

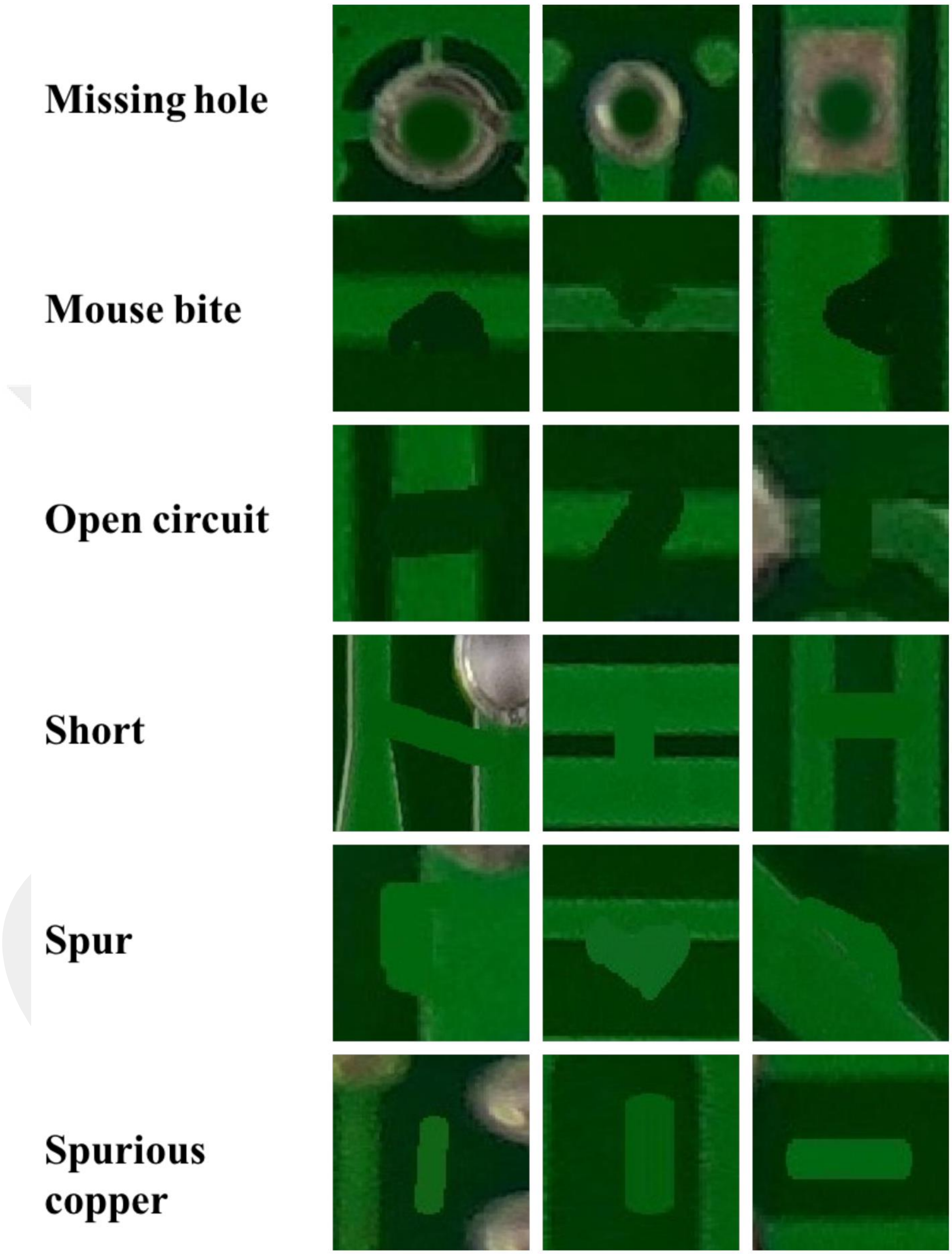

Figure 6 Main Copper Defects (Huang and Wei, 2019)

2.7 Data Collection Procedure

2.7.1 Sampling

The study utilizes a foundational dataset of 690 images, which served as the basis for augmentation. Each augmented dataset was generated to contain approximately 2,000 images, resulting in a total of over 50,000 images. A uniform sampling process was applied to maintain consistency across all augmentations.

2.7.2 Procedure

Dataset Creation and Augmentation

The original dataset was acquired, and an extensive array of augmentation techniques was applied through Roboflow. Following this, the datasets were structured and prepared in alignment with YOLOv8's requirements for experimentation.

Experimental Execution

Experiments were conducted sequentially on a virtual machine, with the necessary folders uploaded to Google Drive. Before initiating each experiment, the model configuration was carefully verified. Upon completion, the results were systematically stored and organized for subsequent analysis.

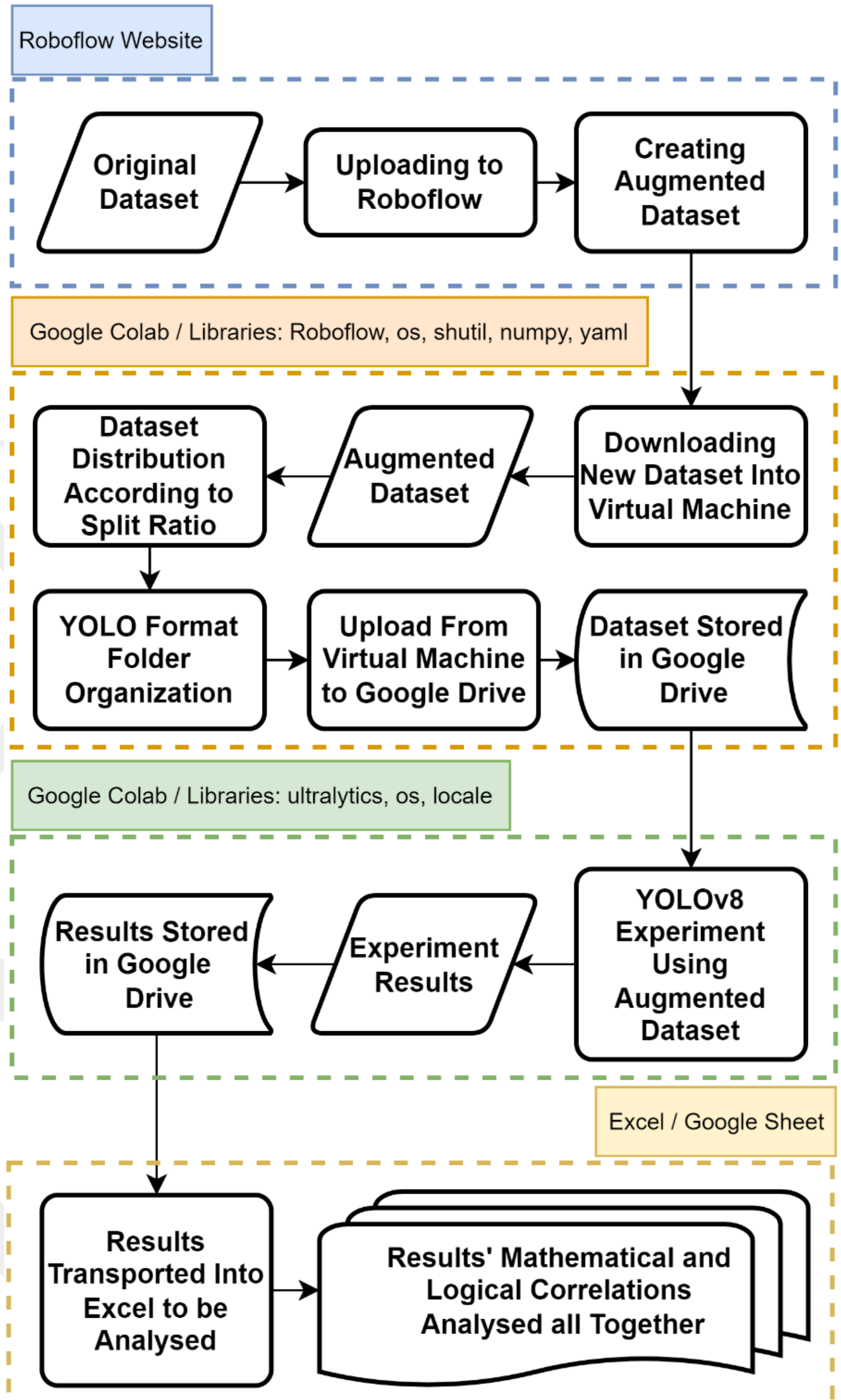

Figure 7 General Setup and Procedure

2.8 Experimental Setup

2.8.1 Control and Treatment Groups

The original dataset functioned as the control group, establishing a baseline for comparison with the augmented datasets. The effectiveness of each augmentation technique was evaluated by comparing its results against this baseline.

2.8.2 Workflow

Baseline Establishment:

Initial experiments were conducted using the original dataset to establish a baseline for subsequent comparisons.

Augmented Dataset Testing:

The augmented datasets were tested using identical procedures, with the results compared both to the baseline and across the various augmentation techniques.

Documentation and Analysis:

Results were recorded in Excel and Google Sheets, and confusion matrices were analyzed to compare performance across different experiments.

2.9 Data Augmentation Techniques

Three primary categories of data augmentation techniques were utilized [6, 7, 8, 9, 10]:

1. Image-Based Augmentations: Techniques such as flipping, rotation, blurring, and color adjustments were applied to enhance model generalization and increase dataset diversity.
2. Bounding Box-Based Augmentations: These methods focused on refining object localization and enhancing the model's capability to detect small, subtle defects. In augmentations involving rotations (ranging from 10 to 45 degrees), the background remained unchanged, except in cases where rotations necessitated blacking out areas previously occupied by the original bounding box.

3. Combination Techniques: Pairing multiple augmentation techniques created complex variations that optimized model performance, thereby enhancing robustness in defect detection.

2.10 Implementation Details

Software and Algorithms:

YOLOv8 and Python were employed to implement the experiments, with Roboflow managing data augmentation and Excel utilized for data analysis.

Parameter Settings:

Each experiment was conducted over 300 epochs with an image resolution of 640x640 pixels. Default parameters were maintained throughout to ensure consistency across all experiments.

Automation:

A four-part script was developed to automate data preparation, ensuring both consistency and efficiency across all datasets.

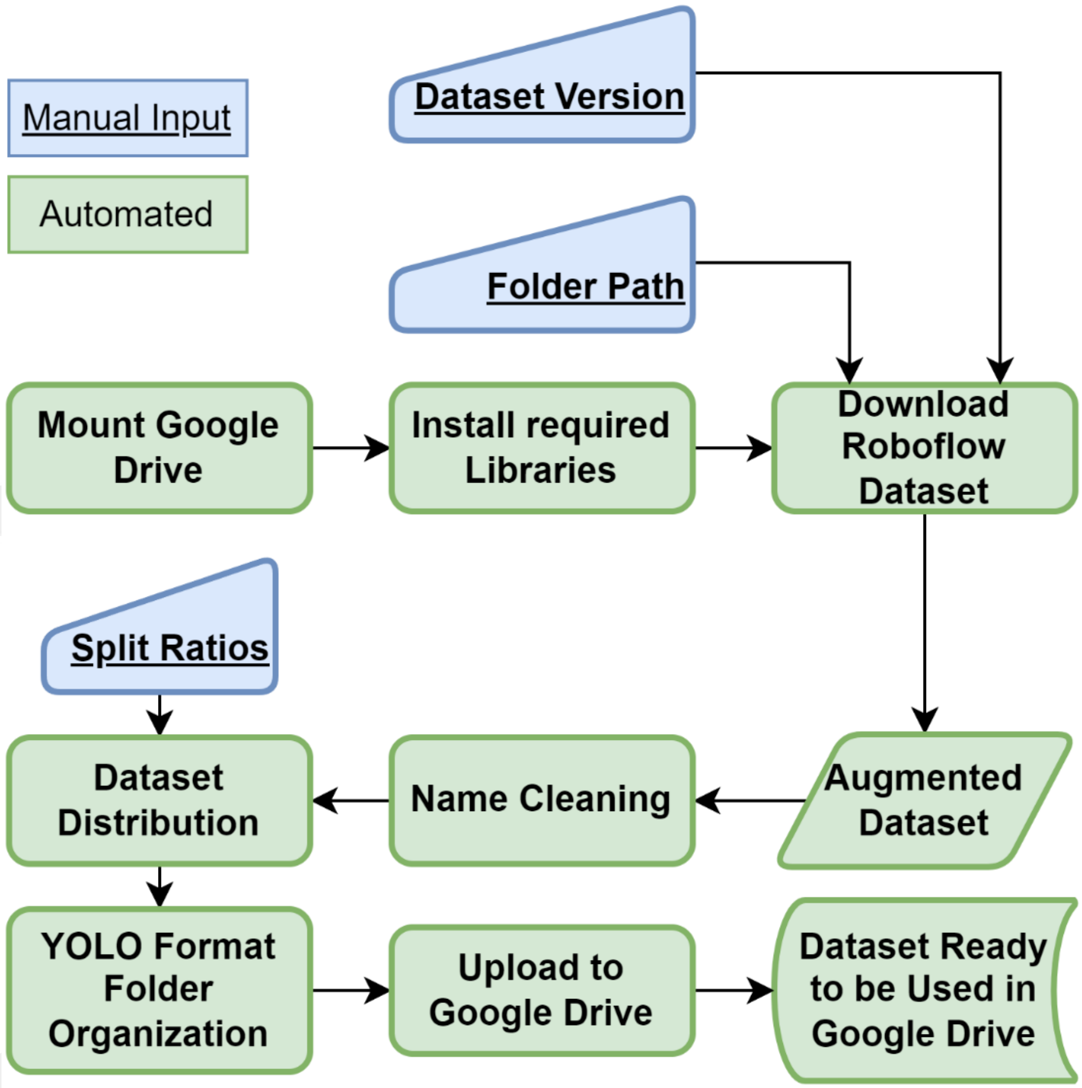

Figure 8 Automated Process

2.11 Data Processing Analysis

Preprocessing:

Images were resized to 640x640 pixels, and key metrics—including precision, recall, mAP50, and F1 score—were calculated. Additionally, analysis included a custom "power score," derived from normalized confusion matrices. This power score is calculated by aggregating normalized scores of each defect category, producing an overall performance score out of 6. It serves as a comprehensive metric that combines multiple evaluation criteria to assess the model's effectiveness across various defect types.

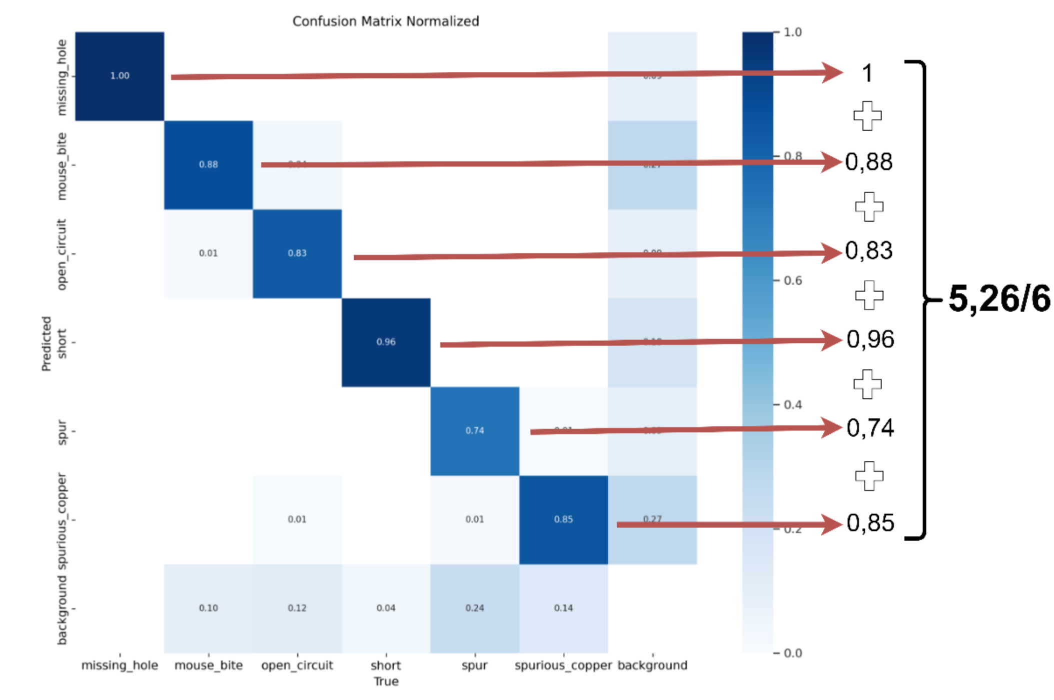

Figure 9 Power Score

2.12 Validation and Testing

The 'Black Edge Fit' preprocessing technique was evaluated across four different train-test-validation split ratios but did not yield the anticipated results. This outcome was unexpected, as the technique is generally recognized for preserving aspect ratios, minimizing information loss, maintaining model consistency at 640x640 resolution, and reducing overfitting. Contrary to expectations, experiments indicated that using no preprocessing led to superior performance. As a result, the two split ratios that achieved the highest power scores—'No Preprocessing 70-15-15' (Power Score: 5.58/6) and 'No Preprocessing 80-10-10' (Power Score: 5.56/6)—were chosen to ensure robust and reliable findings. Metrics such as precision, recall, mAP50, and the power score provided a comprehensive basis for assessing model performance.

2.13 Replication and Reproducibility

This study is fully replicable, as researchers can reproduce the experiments by adhering to the outlined procedures, utilizing the specified dataset and augmentation techniques, and maintaining the same experimental conditions.

2.14 Limitations

The main limitation of this study was the restricted number of images available for augmentation; however, this was addressed by uniformly applying each augmentation technique across all experiments. Minor logistical challenges with Google Colab and Google Drive were managed through custom scripts, ensuring that these issues did not affect the study's outcomes.

CHAPTER 3

RESULTS AND DISCUSSION

In this section, we present and analyze the results from a series of experiments designed to assess the effectiveness of various data augmentation techniques on PCB defect detection using YOLOv8. This study was meticulously structured to enable systematic comparison of different augmentation methods, with the goal of identifying the optimal technique for enhancing detection of small component defects in scenarios with limited datasets. Initially, we examined the effects of varying train-test-validation split ratios without any augmentation to establish a baseline, which allowed us to determine the most effective split ratio to apply in subsequent experiments. The study was organized into several phases: image-based augmentation, bounding box-based augmentation, and combined augmentation techniques. Each phase was rigorously analyzed to uncover mathematical and logical relationships, leading to conclusions regarding the efficacy of each approach. These findings were then synthesized to create a comprehensive overview, providing valuable insights into best practices for data augmentation in PCB defect detection.

3.1 Split Experiments

Table 1 Split Experiment Results

Split (Train - Test - Validation)	Experiment Name	Precision	Recall	mAP50	F1 Score	Power Score #/6
60-20-20	No Preprocessing	0.954	0.881	0.926	0.916	5.28

Table 1 (continued)

60-20-20	Fit (Black Edge)	0.938	0.867	0.913	0.901	5.26
70-15-15	No Preprocessing	0.952	0.926	0.947	0.939	5.58
70-15-15	Fit (Black Edge)	0.949	0.884	0.929	0.916	5.25
70-20-10	No Preprocessing	0.954	0.934	0.948	0.944	5.53
70-20-10	Fit (Black Edge)	0.955	0.883	0.933	0.917	5.33
80-10-10	No Preprocessing	0.963	0.915	0.947	0.938	5.56
80-10-10	Fit (Black Edge)	0.955	0.889	0.922	0.920	5.29

Note. Power score is calculated by adding all the normalized confusion matrix outputs together to get a score out of 6.

3.1.1 Correlation 1: Split Ratio vs. Performance

Observation: Precision, recall, mAP50, and F1 score exhibit variability across different split ratios.

Analysis:

- **Precision & Recall**: With an increase in training data (e.g., shifting from a 60-20-20 to an 80-10-10 split), the "No Preprocessing" approach generally achieves higher precision and recall. This suggests that providing the model with a larger volume of unaltered training images allows it to better learn the true characteristics of the defects.

- **Fit (Black Edge)**: Conversely, in scenarios employing "Fit (Black Edge)," a slight decline in recall is observed as the training data increases. This reduction may imply that the black edges introduce a level of noise or distract from the core features of the images, potentially leading to less effective feature learning by the model.

3.1.2 Correlation 2: Confusion Matrix Score

- **Observation**: The confusion matrix score, rated out of 6, serves as an aggregate performance measure, with higher scores generally indicating greater precision, recall, mAP50, and F1 score.
- **Analysis**: The "No Preprocessing" method consistently achieves higher confusion matrix scores compared to the "Fit (Black Edge)" approach. This indicates that using unaltered images allows the model to better differentiate between defect types, free from the potential interference of black edges. For instance, with a 70-20-10 split, the "No Preprocessing" method attains a score of 5.53, while the "Fit (Black Edge)" method achieves 5.33. This discrepancy reflects a more accurate and consistent performance across defect types with the "No Preprocessing" approach.

3.1.3 Correlation 3: Logical Interpretation

- **Image Preprocessing Impact**: The addition of black edges may influence the model's ability to accurately learn features positioned centrally or near the edges of an image. Given that the YOLOv8 model scans the entire image to detect objects, modifications at the borders—such as black edges—can lead to slight misinterpretations, contributing to the observed performance decline in "Fit (Black Edge)" scenarios.
- **Split Ratios**: A greater proportion of training data generally enhances model performance, as reflected in the metrics. However, this improvement plateaus

when transitioning from a 70-20-10 to an 80-10-10 split, particularly in the "Fit (Black Edge)" case. This trend suggests that increasing training data beyond a certain threshold may yield diminishing returns, especially when preprocessing methods are applied.

3.1.4 Findings

For the forthcoming experiments, we selected the 70-15-15 and 80-10-10 splits with "No Preprocessing" as these configurations strike an effective balance between providing sufficient training data and sustaining high performance across all metrics, particularly the confusion matrix score. The "Fit (Black Edge)" method may be retained for specific cases where preserving the original aspect ratio is essential; however, this approach will be used with caution due to its potential negative impact on model performance.

3.2 Image Based Experiments

Table 2 Image Based 70-15-15 Experiment Results

Experiment Name	**Precision**	**Recall**	**mAP50**	**F1 Score**	**Power Score #/6**
Blur	0.992	0.988	0.993	0.990	5.93
Flip Vertical	0.972	0.965	0.980	0.969	5.78
Flip Vertical Horizontal	0.982	0.968	0.985	0.975	5.85
Grey Scale 25%	0.978	0.950	0.971	0.964	5.71
Grey Scale 100%	0.992	0.976	0.991	0.984	5.90
Hue 25%	0.994	0.987	0.993	0.991	5.93
Hue 100%	0.989	0.990	0.994	0.990	5.96
Noise 1%	0.996	0.985	0.994	0.990	5.92
Noise 2%	0.986	0.978	0.990	0.982	5.88
Random Crop 20%	0.992	0.979	0.991	0.985	5.89

Table 2 (continued)

Random Crop 30%	0.990	0.990	0.994	0.990	5.96
Rotation 10°	0.989	0.985	0.990	0.987	5.93
Rotation 15°	0.991	0.957	0.984	0.974	5.82
Rotation 20°	0.988	0.988	0.992	0.988	5.93
Rotation 25°	0.987	0.976	0.987	0.982	5.89
Rotation 45°	0.978	0.964	0.979	0.971	5.81
Rotation90° CW&CCW	0.975	0.958	0.978	0.967	5.78
Rotation90° CW&CCW&UD	0.981	0.943	0.968	0.962	5.73
Saturation 30%	0.992	0.985	0.992	0.988	5.94
Saturation 70%	0.991	0.981	0.994	0.986	5.91

Table 3 Image Based 80-10-10 Experiment Results

Experiment Name	**Precision**	**Recall**	**mAP50**	**F1 Score**	**Power Score #/6**
Blur	0.993	0.988	0.991	0.990	5.96
Flip Vertical	0.980	0.974	0.985	0.977	5.87
Flip Vertical Horizontal	0.985	0.967	0.985	0.976	5.81
Grey Scale 25%	0.983	0.957	0.976	0.970	5.78

Table 3 (continued)

Grey Scale 100%	0.983	0.972	0.985	0.978	5.89
Hue 25%	0.987	0.994	0.994	0.990	5.98
Hue 100%	0.994	0.985	0.993	0.990	5.92
Noise 1%	0.981	0.990	0.993	0.985	5.95
Noise 2%	0.986	0.983	0.992	0.984	5.90
Random Crop 20%	0.990	0.994	0.994	0.992	5.96
Random Crop 30%	0.989	0.991	0.994	0.990	5.94
Rotation 10°	0.992	0.990	0.995	0.991	5.95
Rotation 15°	0.997	0.982	0.990	0.989	5.92
Rotation 20°	0.995	0.979	0.990	0.987	5.89
Rotation 25°	0.994	0.967	0.990	0.980	5.89
Rotation 45°	0.973	0.972	0.984	0.972	5.84
Rotation90° CW&CCW	0.977	0.964	0.979	0.971	5.79
Rotation90° CW&CCW& UD	0.963	0.956	0.978	0.960	5.79
Saturation 30%	0.989	0.989	0.993	0.989	5.95

Table 3 (continued)

Saturation 70%	0.989	0.989	0.993	0.989	5.93

3.2.1 Correlation 1: Augmentation Techniques vs. Performance

Observation: Across both 70-15-15 and 80-10-10 splits, the metrics of precision, recall, mAP50, and F1 score fluctuate depending on the augmentation technique. While trends are largely consistent across both splits, minor performance variations appear for specific augmentations.

Analysis:

Blur:
Similarity: In both the 70-15-15 and 80-10-10 splits, the Blur augmentation consistently performs well across all metrics. This suggests that small object detection remains effective even with slight blurring, implying that the model emphasizes broader features over fine details.

Flipping (Vertical & Horizontal):
Similarity: In both splits, flipping results in a minor performance drop, particularly in recall, likely due to orientation changes. This indicates that flipping can mislead the model when detecting small objects that typically have a specific orientation.

Greyscale (25% & 100%):
Similarity: In both splits, the 100% greyscale augmentation outperforms the 25% version, suggesting that color information is less essential for detecting small objects. The model appears to prioritize shape and texture, allowing it to maintain performance even when color is entirely removed.

Hue & Saturation:

Similarity: Adjustments in hue and saturation yield similar high performance across both splits. This indicates that variations in color, such as shifts in hue or changes in vibrancy, do not significantly affect small object detection, with the model focusing more on structural features like shape and texture.

Noise (1% & 2%):

Similarity: Adding 1% noise outperforms 2% in both the 70-15-15 and 80-10-10 splits. A minimal level of noise enhances model robustness, whereas higher noise levels obscure essential features, reducing detection accuracy.

Random Crop (20% & 30%):

Difference: In the 70-15-15 split, the 30% crop slightly outperforms the 20% crop, indicating that exposing the model to a broader range of object sizes and positions improves generalization, particularly for small objects. Conversely, in the 80-10-10 split, the 20% crop performs marginally better, suggesting that variability in object positioning is beneficial but that excessive cropping may not yield additional improvements.

Rotation (Various Angles):

Similarity: Smaller rotations (10°-20°) retain high performance in both splits, while larger rotations (25° and 45°) show a performance decline, particularly in recall. This implies that minor rotations aid the model in generalizing across different orientations, but larger rotations distort critical features, negatively impacting detection.

Difference: In the 70-15-15 split, the 15° rotation underperforms compared to other angles due to insufficient variation. In the 80-10-10 split, however, this effect is less noticeable, with the 15° rotation performing similarly to other smaller rotations.

Rotation 90° (CW, CCW, UD):

Similarity: Extreme 90° rotations consistently result in lower performance across both splits, particularly in recall and F1 score. This suggests that the objects in the dataset have a natural orientation, and significant rotation disrupts the model's ability to accurately detect them.

3.2.2 Correlation 2: Confusion Matrix Score

Observation: The Confusion Matrix Score provides an overall measure of performance, with higher scores typically corresponding to better precision, recall, mAP50, and F1 scores in both the 70-15-15 and 80-10-10 splits. While trends remain consistent across both splits, certain augmentations exhibit minor performance variations.

Analysis:

- **Overall Trends**:
  - **Similarity**: In both splits, augmentations that introduce moderate alterations—such as Blur, Noise 1%, Rotation 10°, and Greyscale 100%—consistently achieve higher Confusion Matrix Scores. This indicates more accurate and stable performance across all defect classes, as these moderate augmentations assist the model in generalizing without disrupting essential features critical for detecting small objects.
  - **Difference**: In the 80-10-10 split, the Hue 25% adjustment also contributes positively to the Confusion Matrix Score, suggesting that this subtle hue change works especially well under this split. In contrast, it does not appear as a top performer in the 70-15-15 split.
- **Lower Scores**:
  - **Similarity**: In both splits, techniques that greatly alter the image's orientation or appearance—such as Flipping and Rotation 90°—tend to yield lower Confusion Matrix Scores. While these techniques provide diversity to the training data, they can mislead the model by distorting critical features necessary for accurate small object detection.
  - **Difference**: In the 80-10-10 split, larger rotations (e.g., 25° and 45°) also contribute to lower scores, though the performance drop is less severe than in the 70-15-15 split. This suggests that the 80-10-10 split may offer slightly greater tolerance for these larger rotations, though they still generally degrade overall performance.

3.2.3 Correlation 3: Logical Interpretation

Augmentation Impact:

- **Moderate Augmentations**:
    - **Similarity**: Across both the 70-15-15 and 80-10-10 splits, moderate augmentations—such as blur, small rotations, and mild noise—consistently improve the model’s robustness. These augmentations introduce subtle variations without substantially altering the objects' natural appearance, preserving the essential features required for accurate detection. Consequently, they maintain high accuracy and consistency in detecting small objects.
- **Extreme Augmentations**:
    - **Similarity**: In both splits, extreme augmentations that introduce significant changes—such as large rotations, full flips, and heavy noise—tend to impair the model's accuracy in detecting small objects. These techniques often distort image features excessively, deviating too far from the model's learned representations, which can negatively impact detection performance.

Small Object Detection:

- **Shape and Texture**:
    - **Similarity**: In both experiments, the strong performance of greyscale, hue, and saturation adjustments suggests that small object detection relies more on structural attributes, such as shape and texture, rather than on color. The model's high performance, even when color information is reduced or altered, further supports this finding.
- **Orientation Sensitivity**:
    - **Similarity**: Both splits demonstrate that small objects are likely sensitive to orientation. Techniques involving large rotations or flips generally lead to reduced performance, indicating that drastic orientation changes can hinder the model’s ability to accurately detect small objects.

3.2.4 Findings

For small object detection, subtle augmentations such as Blur, Noise at 1%, Rotation at 10°, and Hue Adjustments consistently yield the best results across both the 70-15-15 and 80-10-10 splits. These techniques enhance the model's robustness by introducing minor adjustments that preserve the essential features crucial for accurate detection. Additionally, Greyscale and Random Crop augmentations also demonstrate strong performance, with the 70-15-15 split showing a preference for a 30% crop and the 80-10-10 split favoring a 20% crop, indicating that slight variations in cropping percentages can influence detection effectiveness.

Conversely, extreme augmentations—such as large rotations and full flips—should be applied cautiously, as they can distort the features essential for small object detection. These transformations pose challenges by shifting objects significantly from their natural orientations or distorting their appearance, ultimately leading to reduced detection accuracy across both splits.

3.3 Bounding Box Experiments

Table 4 Bounding Box 70-15-15 Experiment Results

Experiment Name	**Precision**	**Recall**	**mAP50**	**F1 Score**	**Power Score #/6**
Blur	0.986	0.986	0.991	0.986	5.91
Flip Vertical	0.985	0.985	0.992	0.985	5.92
Flip Vertical Horizontal	0.990	0.978	0.992	0.984	5.90
Noise 1%	0.988	0.994	0.993	0.991	5.98
Noise 2%	0.994	0.987	0.995	0.991	5.96
Rotation 10°	0.990	0.984	0.991	0.987	5.90

Table 4 (continued)

Rotation 15°	0.994	0.993	0.995	0.994	5.96
Rotation 20°	0.982	0.991	0.993	0.986	5.96
Rotation 25°	0.988	0.984	0.993	0.986	5.90
Rotation 45°	0.973	0.981	0.991	0.977	5.89
Rotation90° CW&CCW	0.977	0.966	0.983	0.972	5.81
Rotation90° CW&CCW&UD	0.973	0.970	0.988	0.972	5.82

Table 5 Bounding Box 80-10-10 Experiment Results

Experiment Name	**Precision**	**Recall**	**mAP50**	**F1 Score**	**Power Score #/6**
Blur	0.988	0.984	0.987	0.986	5.93
Flip Vertical	0.991	0.989	0.991	0.990	5.93
Flip Vertical Horizontal	0.988	0.988	0.991	0.988	5.92
Noise 1%	0.989	0.990	0.993	0.989	5.94
Noise 2%	0.987	0.986	0.994	0.987	5.96
Rotation 10°	0.990	0.988	0.991	0.989	5.96
Rotation 15°	0.991	0.995	0.994	0.993	5.98
Rotation 20°	0.987	0.992	0.993	0.990	5.93

Table 5 (continued)

Rotation 25°	0.992	0.992	0.995	0.992	5.98
Rotation 45°	0.990	0.988	0.993	0.989	5.94
Rotation90° CW&CCW	0.981	0.990	0.994	0.986	5.93
Rotation90° CW&CCW&UD	0.984	0.979	0.993	0.982	5.87

3.3.1 Correlation 1: Augmentation Techniques vs. Performance

Observation: In both the 70-15-15 and 80-10-10 splits, precision, recall, mAP50, F1 score, and Confusion Matrix Scores vary based on different bounding box augmentation techniques. While certain trends remain consistent across both splits, some minor variations are observed.

Analysis:

Blur:
Similarity: In both the 70-15-15 and 80-10-10 splits, applying blur within the bounding box consistently results in strong performance across all metrics. This suggests that blurring the region of interest within the bounding box does not obscure critical features, allowing the model to accurately detect small objects.

Flipping (Vertical & Horizontal):
Similarity: Flipping within the bounding box performs well in both splits, with Vertical Flip consistently outperforming Vertical Horizontal Flip. A slight drop in recall is observed with the latter, suggesting that while the model generalizes well with simple flips, more complex flips (like vertical-horizontal) may pose minor challenges, particularly for small objects.

Difference: The 80-10-10 split shows particularly high performance with vertical flips, slightly outperforming the 70-15-15 split, suggesting better generalization to orientation changes in this configuration.

Noise (1% & 2%):

Similarity: Adding noise within the bounding box proves effective in both splits, with 2% noise showing peak performance in the 70-15-15 split, while 1% noise performs slightly better in the 80-10-10 split. This implies that introducing noise inside the bounding box increases model robustness against minor imperfections and real-world distortions.

Rotation (Various Angles):

Similarity: Small rotations (10°-20°) perform well in both splits, while larger rotations (45°) show greater variability in performance.

Difference: In the 70-15-15 split, larger rotations like 45° lead to a decrease in detection accuracy, indicating that such transformations misalign object features, making detection more challenging. In contrast, in the 80-10-10 split, even the 45° rotation maintains high performance, suggesting that the model in this configuration adapts more effectively to these transformations.

Rotation 90° (CW, CCW, UD):

Similarity: Extreme 90° rotations consistently result in the lowest performance across both splits, especially in recall and F1 score. This indicates that significant orientation changes within the bounding box distort object features excessively, reducing detection accuracy.

3.3.2 Correlation 2: Confusion Matrix Score

Observation: The Confusion Matrix Score serves as a comprehensive indicator of overall performance, with higher scores generally aligned with improved precision, recall, mAP50, and F1 scores. Certain consistent trends emerge across the 70-15-15 and 80-10-10 splits, with specific augmentation techniques producing better scores.

Analysis:

- **Overall Trends**:
  - **Similarity**: Techniques such as Noise 1%, Rotation 15°, and Rotation 20° (for the 70-15-15 split) or Rotation 25° (for the 80-10-10 split) consistently achieve the highest Confusion Matrix Scores. These augmentations strike a balance between introducing variability and preserving essential object features, resulting in consistent performance across all defect categories.
  - **Difference**: The 80-10-10 split demonstrates slightly better performance with Rotation 25°, suggesting that this configuration tolerates slightly larger rotations while maintaining high detection accuracy, whereas Rotation 20° is more effective in the 70-15-15 split.
- **Lower Scores**:
  - **Similarity**: In both splits, augmentations that significantly alter the bounding box, such as Rotation 45° and Rotation 90° (CW, CCW, UD), yield lower Confusion Matrix Scores. These extreme augmentations likely introduce excessive variation within small object features, resulting in reduced detection accuracy.
  - **Difference**: In the 80-10-10 split, the performance drop for extreme rotations (like 90°) is less pronounced compared to the 70-15-15 split, suggesting that the model might better accommodate extreme augmentations under this configuration. However, these techniques still underperform overall, providing a baseline for future experiments.

3.3.3 Correlation 3: Logical Interpretation

Augmentation Impact:

- **Moderate Augmentations**:
  - **Similarity**: In both splits, moderate augmentations, such as slight noise and small rotations within the bounding box, enhance the model's robustness. These techniques introduce subtle changes without distorting essential features, thereby maintaining high detection accuracy and consistency across key metrics.

    - **Difference**: In the 80-10-10 split, moderate rotations (up to 25°) contribute to improved robustness compared to the 70-15-15 split, suggesting that this configuration can handle slightly larger rotations without compromising detection accuracy.
- **Extreme Augmentations**:
    - **Similarity**: In both splits, extreme augmentations—such as 90° rotations or large flips—distort small object features excessively. These transformations interfere with the model's ability to accurately detect the object, leading to a performance drop.

Small Object Detection:

- **Focus on Core Features**:
    - **Similarity**: Across both splits, the strong performance achieved with minor noise and small rotations suggests that small object detection benefits from augmentations that slightly modify core features within the bounding box. These subtle changes help the model generalize to real-world scenarios where minor variations are common.
- **Orientation Sensitivity**:
    - **Similarity**: The reduced performance observed with 90° rotations indicates that small objects have a preferred orientation, even within the bounding box. Drastic changes in orientation, such as extreme rotations, disrupt the model's ability to accurately detect small objects in both splits.

3.3.4 Findings

For bounding box augmentations, the most effective techniques across both the 70-15-15 and 80-10-10 splits are those introducing subtle to moderate variations, including Noise at 1%, Rotation between 15°-20° (or up to 25° in the 80-10-10 split), and Vertical Flip (notably in the 80-10-10 split). These augmentations enhance the model's robustness by making minor adjustments to object features within the bounding box without distorting them, resulting in consistent and accurate detection of small objects.

In contrast, extreme augmentations—such as large rotations and 90° flips—consistently underperform across both splits. These transformations excessively distort essential object features, impairing the model’s ability to accurately detect small objects. Consequently, such augmentations should be applied with caution, particularly in applications demanding high detection accuracy.

3.4 Generalization of Non-Paired Experiment Correlations

3.4.1 Mathematical and Logical Correlations

3.4.1.1 Performance Across Different Augmentation Techniques

Blur:

- **Mathematical Correlation**: Blur consistently achieves high precision, recall, mAP50, and F1 scores across both image-based and bounding box augmentations in both split ratios. However, the slight decrease in performance in bounding box experiments suggests that blurring the entire image is more effective than restricting it to the object within the bounding box.
- **Logical Interpretation**: The model's robustness in detecting small objects even when the image is blurred indicates that detection is more reliant on overall shape and context than on fine details. Blurring within the bounding box might obscure critical texture details essential for distinguishing small objects.

Flipping (Vertical & Horizontal):

- **Mathematical Correlation**: Flipping tends to result in slightly lower recall and Confusion Matrix Scores, especially in bounding box experiments compared to full image augmentation, with this trend consistent across both split ratios.
- **Logical Interpretation**: Changing object orientation through flipping can introduce inconsistencies in the model's learning process. When only the bounding box is flipped, the model may struggle to reconcile the orientation of the object with its surrounding context, resulting in marginally reduced performance.

Noise:

- **Mathematical Correlation**: Noise, particularly at 1%, consistently improves model robustness across all configurations, with optimal performance in the bounding box 80-10-10 setup. This suggests that introducing noise within the object region enhances the model’s ability to generalize.
- **Logical Interpretation**: Adding noise mimics real-world imperfections, thereby increasing the model's resilience. Noise within the bounding box is especially effective as it directly impacts on the object’s features, making the model less sensitive to minor distortions or artifacts.

Rotation (Various Angles):

- **Mathematical Correlation**: Small rotations (10° to 25°) generally lead to high performance, whereas larger rotations (45° and 90°) reduce accuracy, especially in bounding box experiments.
- **Logical Interpretation**: Small rotations help the model learn to detect objects at slightly varied orientations, supporting generalization. However, larger rotations disrupt feature alignment, particularly when limited to the bounding box, causing confusion and diminishing detection accuracy.

Hue & Saturation (from earlier image-based experiments):

- **Mathematical Correlation**: Adjustments in hue and saturation, particularly in image-based augmentations, demonstrate that the model is not overly dependent on color, as performance remains high. This indicates that the model prioritizes structural features over color.
- **Logical Interpretation**: Changes in hue and saturation are effective in preventing the model from memorizing specific color schemes, enhancing its ability to generalize. This is important for detection under varied lighting conditions and color environments, ensuring robust performance across different scenarios.

3.4.1.2 Performance Across Different Split Ratios

70-15-15 vs. 80-10-10:

- **Mathematical Correlation**: The 80-10-10 split typically yields slightly higher performance metrics than the 70-15-15 split, especially in precision and recall. This trend is consistent across both image-based and bounding box augmentations.
- **Logical Interpretation**: A larger training set, as in the 80-10-10 split, provides the model with more diverse features to learn from, leading to improved generalization and accuracy. However, the marginal improvement suggests that additional data beyond a certain point yield diminishing returns.

3.4.2 Generalized Augmentation for Small Object Detection

Balanced Augmentation for Small Object Detection:

- Augmentations introducing moderate changes—such as slight rotations, noise, and blurring—consistently enhance model performance across both image-based and bounding box methods. Extreme augmentations, such as large rotations or heavy noise, often degrade performance, especially when limited to the bounding box.

Image-Based Augmentation Provides Broader Context:

- Image-based augmentations tend to yield slightly better performance than bounding box augmentations, especially for tasks involving object orientation and contextual learning. This suggests that the broader perspective provided by the entire image is advantageous for small object detection.

Bounding Box Augmentation Enhances Feature Sensitivity:

- Bounding box augmentations direct the model's focus onto critical features within the object area, enhancing sensitivity to small details. However, this approach risks losing contextual information, potentially reducing performance if the augmentation is too extreme.

Smaller Rotations Improve Generalization; Larger Rotations Introduce Ambiguities:

- Small rotations (10° to 25°) are effective for generalizing to slightly varied object orientations, while larger rotations (45° and 90°) introduce ambiguities, leading to decreased performance, particularly in bounding box scenarios.

Noise Augmentation is Effective but Requires Calibration:

- Adding noise, especially within the bounding box, increases robustness to real-world imperfections. However, the noise level must be carefully calibrated, as excessive noise can obscure essential features and reduce detection accuracy.

Higher Training Data Proportions Yield Marginal Gains:

- Increasing the training data proportion from 70-15-15 to 80-10-10 generally improves performance, but the gains are minor. This suggests that while more data is beneficial, the model eventually encounters diminishing returns where additional data has a limited impact on performance.

3.4.3 Findings

Synthesizing insights from all four experimental setups reveals that effective small object detection hinges on a balanced approach to augmentation. This balance maintains essential features while introducing sufficient variability to enhance generalization. Image-based augmentations contribute valuable context, while bounding box augmentations refine the model's sensitivity to specific object features. Split ratios have a modest influence, with the 80-10-10 split providing slight performance gains due to the increased volume of training data.

3.5 Image Based Paired Experiments

In this chapter, we examine the results of pairing different augmentation techniques. Notably, image-based and bounding box-based augmentations were not mixed in any experiment; rather, image-based augmentations were paired solely with other image-based techniques, and bounding box-based augmentations were similarly paired within their category. This strategy allowed for a focused analysis of the impact of each augmentation type on model performance.

For these paired experiments, the 80-10-10 split ratio was selected due to its slightly superior performance in previous tests, as reflected in the average Power Scores across experiments:

- **Average Image-Based 70-15-15 score**: 5.873
- **Average Image-Based 80-10-10 score**: 5.896
- **Average Bounding Box 70-15-15 score**: 5.909
- **Average Bounding Box 80-10-10 score**: 5.939

These results underscore the consistent advantage of the 80-10-10 split.

In this study, augmentation techniques were paired through two distinct methodologies to ensure a non-random selection process:

1. **Geometrically Similar Pairing**: This approach involved pairing augmentations with similar geometric effects to create a dataset with sufficient consistency for the model to deeply learn defect patterns while remaining robust to minor variations. By using comparable augmentations, the model could fine-tune its detection of subtle defects without being overwhelmed by excessive variability.
2. **Power Score-Based Pairing**: Here, one low-performing and one high-performing augmentation were paired to observe their combined impact on model performance. This approach was designed to evaluate whether the strengths of the higher-performing augmentation could compensate for the weaknesses of the lower-performing one and to understand how their interaction influenced overall detection accuracy.

Table 6: Image Based Paired 80-10-10 Experiment Results

Experiment Name	**Precision**	**Recall**	**mAP50**	**F1 Score**	**Power Score #/6**
IB: Blur & Noise 1%	0.991	0.984	0.990	0.988	5.92
IB: CW CCW UD & Noise 1%	0.974	0.963	0.988	0.969	5.79
IB: Flip Vertical & CW CCW	0.976	0.955	0.977	0.965	5.74
IB: Flip Vertical & Hue 25%	0.990	0.978	0.990	0.984	5.90
IB: Flip Vertical & Rotation 10°	0.988	0.983	0.992	0.986	5.92
IB: Hue 25% & Saturation 30%	0.991	0.992	0.991	0.991	5.97
IB: Rotation 45° & Random Crop 20%	0.972	0.963	0.983	0.968	5.79

Table 7: Image Based Lone vs. Paired Experiment Results

Lone 1	**Power Score #/6**	**Lone 2**	**Power Score #/6**	**Paired**	**Power Score #/6**
Only Blur	5.96	Only Noise1%	5.95	IB: Blur & Noise 1%	5.92

Table 7 (continued)

Only CW CCW UD	5.79	Only Noise 1%	5.95	IB: CW CCW UD & Noise 1%	5.79
Only Flip Vertical	5.87	Only CW CCW	5.79	IB: Flip Vertical & CW CCW	5.74
Only Flip Vertical	5.87	Only Hue 25%	5.98	IB: Flip Vertical & Hue 25%	5.90
Only Flip Vertical	5.87	Only Rotation 10°	5.95	IB: Flip Vertical & Rotation 10°	5.92
Only Hue	5.98	Only Saturation 30%	5.95	IB: Hue 25% & Saturation 30%	5.97
Only Rotation 45°	5.84	Only Random Crop 20%	5.96	IB: Rotation 45° & Random Crop 20%	5.79

3.5.1 Correlation 1: Analysis of Individual vs. Paired Augmentations

Blur & Noise 1%:

- **Change in Performance**: A slight decrease in performance occurred when these augmentations were combined, with scores dropping from 5.96 and 5.95 individually to 5.92 in pairing.

- **Reasoning**: Blur and noise both introduce forms of distortion—blur affects overall clarity, while noise adds pixel-level randomness. Individually, each can enhance robustness, but together they might introduce excessive distortion, potentially obscuring subtle features. This combination could overwhelm the model, making it difficult to differentiate critical details from random noise.

CW CCW UD & Noise 1%:

- **Change in Performance**: A notable drop in performance was observed, with scores falling from 5.95 for Noise 1% alone to 5.79 when paired with CW CCW UD.
- **Reasoning**: CW CCW UD introduces extreme orientation changes, which already challenge the model. Adding noise doesn't mitigate the disorienting effects of these rotations, leading to performance similar to CW CCW UD alone. The combination might overwhelm the model as it struggles to align random noise with drastic orientation shifts.

Flip Vertical & CW CCW:

- **Change in Performance**: Performance decreased slightly when paired (5.74 vs. 5.87 and 5.79).
- **Reasoning**: Flip Vertical and CW CCW both alter orientation in different ways, introducing extensive variability that may confuse the model. The need to adapt to multiple, conflicting orientation changes likely challenges the model, reducing detection accuracy.

Flip Vertical & Hue 25%:

- **Change in Performance**: This pairing led to a slight performance decrease compared to Hue 25% alone (5.98 to 5.90), but was still better than Flip Vertical alone.
- **Reasoning**: Hue 25% adds beneficial color variation, preventing the model from over-relying on specific colors. However, pairing with Flip Vertical, which adds

orientation variability, slightly offsets these gains, suggesting that while color variation is helpful, the orientation change may counteract some of its advantages.

Flip Vertical & Rotation 10°:

- **Change in Performance**: Performance improved slightly to 5.92 compared to Flip Vertical alone but decreased from Rotation 10° alone.
- **Reasoning**: Rotation 10° introduces a controlled orientation change that complements Flip Vertical, creating moderate variability without overwhelming the model. The slight drop from 5.95 to 5.92 suggests that, while beneficial, the added flip variability introduces a minor challenge to the model.

Hue 25% & Saturation 30%:

- **Change in Performance**: This combination performed slightly worse than Hue 25% alone (5.98 to 5.97), though better than Saturation 30% alone.
- **Reasoning**: Hue and saturation adjustments affect color differently—hue changes color, while saturation changes intensity. Together, they create a model that's robust against color shifts, though the minor drop suggests that the saturation adjustment might slightly dampen the benefits of hue adjustment.

Rotation 45° & Random Crop 20%:

- **Change in Performance**: A performance decrease was observed compared to Random Crop 20% alone.
- **Reasoning**: Rotation 45° introduces significant orientation changes, making object detection more challenging. When paired with Random Crop 20%, which varies object positioning and size, the combination intensifies these challenges, reducing performance as the object becomes both rotated and partially obscured, complicating accurate detection.

3.5.2 Correlation 2: Generalized Insight from Paired vs. Individual Augmentations

- **Combining Similar Types of Augmentations Can Lead to Diminished Returns**: Pairing augmentations with similar effects, such as Blur with Noise—both of which introduce distortions—slightly reduces performance. Likewise, combining CW CCW UD with Noise 1% does not enhance performance, as the extreme rotations already challenge the model, and additional noise only increases complexity without adding robustness.
- **Orientation Changes Are Best Applied in Moderation**: Pairings involving significant orientation changes, like Flip Vertical with CW CCW, tend to reduce performance. The model achieves better results when orientation adjustments are applied more conservatively, such as pairing Flip Vertical with a small rotation (10°), which introduces manageable variability.
- **Color Adjustments Are Effective and Mostly Complementary**: Pairing color adjustments, such as Hue and Saturation, generally enhances performance by helping the model generalize across various lighting and color environments. However, saturation adjustments need to be balanced carefully to avoid slightly diminishing the benefits of hue adjustments.
- **Pairing Should Consider the Nature of the Augmentations**: Successful pairings, like Flip Vertical with Hue 25%, work because they address different aspects of the model's learning—orientation and color variation. In contrast, less effective pairings, such as CW CCW UD with Noise 1%, combine challenges that are either too similar or too extreme, overwhelming the model and leading to decreased performance.

3.5.3 Findings

A comparison between individual augmentations and their paired counterparts reveals that not all augmentations perform better in combination. The effectiveness of a pairing largely hinges on whether the augmentations complement one another or introduce excessive variability that the model struggles to manage.

When designing augmentation strategies for small object detection in PCB inspections, it is essential to assess the impact of each augmentation individually and in combination with others. Effective pairings should aim to balance different types of variability—such as orientation, color, and noise—without overwhelming the model with excessive changes, thus maintaining a controlled environment for optimal detection accuracy.

3.6 Bounding Box Paired Experiments

Table 8 Bounding Box Paired 80-10-10 Experiment Results

Experiment Name	Precision	Recall	mAP50	F1 Score	Power Score #/6
BB: Blur & Noise 1%	0.985	0.991	0.993	0.988	5.94
BB: CW CCW UD & Noise 1%	0.979	0.978	0.987	0.979	5.84
BB: Flip Vertical & CW CCW	0.976	0.966	0.983	0.971	5.78
BB: Flip Vertical & Rotation 15°	0.988	0.980	0.993	0.984	5.87

Table 9 Bounding Box Lone vs. Paired Experiment Results

Lone 1	Power Score #/6	Lone 2	Power Score #/6	Paired	Power Score #/6
Only Blur	5.93	Only Noise 1%	5.94	BB: Blur & Noise 1%	5.94
Only CW CCW UD	5.87	Only Noise 1%	5.94	BB: CW CCW UD & Noise 1%	5.84
Only Flip Vertical	5.93	Only CW CCW	5.93	BB: Flip Vertical & CW CCW	5.78
Only Flip Vertical	5.93	Only Rotation 15°	5.98	BB: Flip Vertical & Rotation 15°	5.87

3.6.1 Correlation1: Analysis of Individual vs. Paired Augmentation

BB: Blur & Noise 1%

- **Change in Performance**: Performance slightly decreased when these augmentations were paired, dropping from 5.96 and 5.95 individually to 5.92 when combined.
- **Reasoning**: Blur and noise both introduce forms of distortion—blur impacts overall image clarity, while noise adds pixel-level randomness. Although each augmentation improves robustness individually, their combination may introduce excessive distortion, potentially causing the model to overlook subtle features. This pairing might overwhelm the model by blending important details with random noise.

BB: CW CCW UD & Noise 1%

- **Change in Performance**: A slight performance decrease was observed, with scores dropping from 5.94 and 5.87 individually to 5.84 when paired.
- **Reasoning**: CW CCW UD introduces extreme orientation changes, which already challenge the model. Adding noise compounds this challenge by introducing additional variability that can obscure critical features within the rotated bounding box. This combination likely overwhelms the model, resulting in a minor reduction in classification accuracy.

BB: Flip Vertical & CW CCW

- **Change in Performance**: A significant drop in performance occurred when these augmentations were paired, with scores decreasing from 5.93 individually to 5.78 when combined.
- **Reasoning**: Flip Vertical and CW CCW both alter object orientation in distinct ways, and their combination likely introduces excessive variability, making it challenging for the model to maintain consistent accuracy. These drastic orientation shifts disrupt the model's ability to recognize key features, leading to a more substantial decrease in performance compared to other pairings.

BB: Flip Vertical & Rotation 15°

- **Change in Performance**: Performance slightly decreased to 5.87 when paired, slightly lower than Rotation 15° alone but higher than Flip Vertical alone.
- **Reasoning**: While Rotation 15° introduces beneficial variability on its own, pairing it with Flip Vertical adds more orientation changes than the model can effectively manage. The performance decrease suggests that the combined variability slightly hinders the model's ability to generalize, even though each augmentation was beneficial individually.

3.6.2 Correlation 2: Generalization Insight from Bounding Box Pairings

- **Paired Augmentations Often Lead to Compounded Complexity**: Pairing augmentations that introduce significant changes, especially in orientation (e.g., Flip Vertical & CW CCW), tends to compound complexity for the model, leading to a more pronounced performance drop. The model may struggle to reconcile multiple, conflicting orientation changes, which affects its ability to accurately detect defects within the bounding box.
- **Combining Blur and Noise Maintains Robustness**: The Blur and Noise 1% pairing did not significantly alter performance compared to the individual augmentations, indicating that while the combination does not provide additional benefits, it also does not introduce adverse effects. This suggests a balanced approach to handling various image distortions without compromising model performance.
- **Moderate Rotation is Effective but Requires Caution When Paired**: Rotation 15° was highly effective alone; however, pairing it with Flip Vertical led to a slight performance decline. This implies that while moderate rotations are beneficial, they should be paired cautiously to avoid excessive variability that could impact the model's performance.
- **Noise Can Be Beneficial but Should Not Overwhelm the Model**: Adding noise to already challenging augmentations, such as extreme rotations, can reduce performance, as observed in the CW CCW UD & Noise 1% pairing. Noise should be applied judiciously, particularly when combined with other complex augmentations, to prevent overwhelming the model's capacity for accurate defect detection.

3.6.3 Findings

In the context of bounding box augmentations, it is essential to carefully assess the interactions between different types of augmentations. While some, such as Blur and Noise, can be combined without significantly affecting performance, others—especially those that involve multiple orientation changes—can lead to reduced accuracy. The primary takeaway is that augmentations should be paired to introduce variability while allowing the model to generalize effectively without being overloaded by excessive changes. Balancing augmentation combinations is crucial for maintaining high detection accuracy.

3.7 Generalization of Paired Experiment Correlations

3.7.1 Generalized Correlation Analysis

Performance Trends in Image-Based Pairings:

- **Mathematical Correlation**: In image-based augmentations, pairings that incorporate moderate variability—such as Hue 25% & Saturation 30% or Flip Vertical & Hue 25%—typically yield high performance. Conversely, when augmentations that introduce significant orientation changes are combined (e.g., CW CCW UD & Noise 1%), a noticeable decline in performance occurs.
- **Logical Interpretation**: The preservation of the entire image context in image-based augmentations aids in maintaining overall spatial relationships within the data. This context is beneficial for the model, particularly when detecting small objects, as it allows the model to utilize surrounding information for accurate detection. However, significant alterations to this context or excessive variability can hinder the model's ability to generalize effectively.

Performance Trends in Bounding Box Pairings:

- **Mathematical Correlation**: In bounding box augmentations, pairings that focus on localized changes—such as Blur & Noise 1%—consistently maintain performance. In contrast, pairings that involve multiple orientation changes within the bounding box (e.g., Flip Vertical & CW CCW) lead to a more pronounced decrease in performance compared to their individual effects.
- **Logical Interpretation**: Bounding box augmentations isolate the object of interest, making the model particularly sensitive to changes within this confined region. When paired augmentations drastically alter the object's orientation or introduce excessive variability, the model's ability to recognize the object can be

compromised. This sensitivity is critical for small objects, where even minor misalignments or distortions can significantly impede detection.

Comparative Analysis of Image-Based vs. Bounding Box Pairings:

- **Mathematical Correlation**: In both image-based and bounding box approaches, pairings that involve excessive orientation changes or complex augmentations generally result in decreased performance. However, the negative impact is more pronounced in bounding box augmentations, where the model struggles more with localized changes due to the lack of broader contextual information from the full image.
- **Logical Interpretation**: The broader context provided by image-based augmentations appears to alleviate some challenges associated with paired augmentations. In contrast, while the bounding box approach can be effective in specific scenarios, it renders the model more susceptible to confusion when multiple complex augmentations are applied concurrently.

3.7.2 Generalization about Augmentation Pairing for Small Object Detection

- **Context Preservation is Crucial for Effective Pairing**: Augmentations that maintain or only moderately alter the image context—such as color adjustments like Hue and Saturation—tend to pair effectively, as they enable the model to leverage both the object's features and its surroundings. In contrast, pairings that disrupt this context, particularly in bounding box augmentations, often result in decreased performance.
- **Moderate Variability Enhances Robustness Without Overwhelming the Model**: Pairings that introduce moderate variability—such as small rotations combined with a flip or color adjustments—generally enhance the model's robustness without overwhelming it. These combinations assist the model in generalizing across slightly different scenarios while avoiding excessive complexity.

- **Excessive Orientation Changes Should Be Avoided in Pairings**: Pairings that involve multiple or extreme orientation changes—such as combining CW CCW rotations with vertical flips—tend to confuse the model, leading to a significant drop in performance. Such pairings should be avoided, especially in bounding box augmentations, where accurate detection relies heavily on the object's orientation.
- **Localized Changes in Bounding Box Augmentations Require Careful Balancing**: In bounding box augmentations, where the focus is primarily on the object itself, pairing augmentations should be approached with caution. Combining techniques like Blur and Noise, which target different aspects of the object, can maintain or even slightly enhance performance. However, pairings that focus on similar features or introduce conflicting changes may degrade performance.
- **Pairing Should Complement, Not Complicate, the Model's Learning Process**: Effective augmentation pairings should complement one another by addressing different aspects of the model's learning process. For example, pairing a spatial augmentation like rotation with a color-based augmentation such as hue adjustment can create a balanced approach to variability. Conversely, combinations that complicate the model's learning—by introducing excessive variability, such as pairing significant orientation changes with pixel-level noise—can overwhelm the model and diminish its accuracy.

3.7.3 Findings

When developing augmentation strategies for small object detection, particularly in applications such as PCB defect detection, it is vital to consider the interactions between different augmentations. The objective is to formulate pairings that introduce sufficient variability to enhance robustness without overwhelming the model with excessive alterations. For image-based augmentations, preserving context is essential to ensure that the model retains valuable spatial relationships. In contrast, for bounding box augmentations, the focus should be on localized changes that enhance the model's ability to detect small objects rather than introduce confusion. This thoughtful approach is key to optimizing model performance and accuracy.

3.8 Data Augmentation Order and Its Impact

Throughout this thesis, the dataset augmentation process has been conducted prior to splitting the data into train, test, and validation folders. This approach ensures that every image used during the YOLO experiments is augmented, thereby providing consistent exposure to variations in the data across all stages of model training and evaluation. To assess the efficacy of this methodology, a comparative experiment was conducted where the dataset was first split into train, test, and validation folders, and only the training folder was augmented afterward. The results from this alternative approach were significantly inferior, highlighting the superiority of augmenting the dataset before splitting.

To illustrate the impact of these strategies, a normalized confusion matrix is presented, offering a detailed visualization of the differences in model performance under the two conditions. This analysis underscores the critical importance of consistent augmentation across the dataset to enhance model generalization and performance.

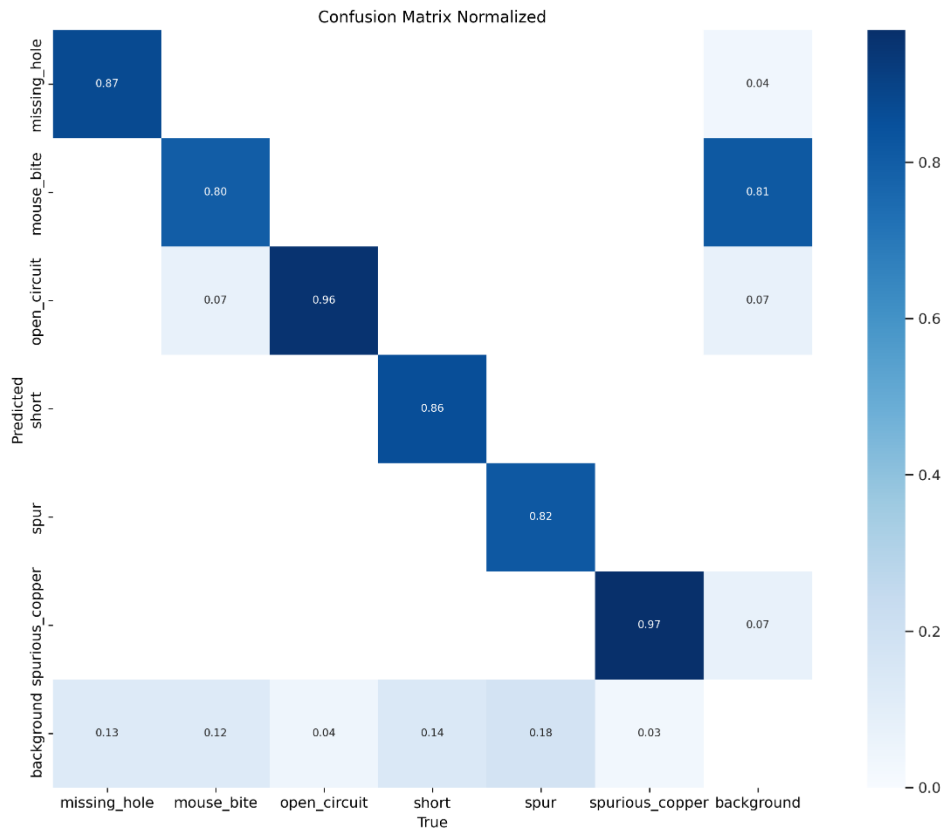

Figure 10 Augmentation Order Changed Experiment Matrix

A potential concern in augmentation-heavy approaches is the risk of overfitting. However, overfitting was effectively mitigated in this study due to two key factors:

YOLO's Training Stopping Criterion: The YOLO framework implements a mechanism to halt training after 100 epochs if there is minimal improvement in key metrics, ensuring the model does not continue to train on redundant patterns.

Training from Scratch: For each experiment, the pre-trained weight document was re-downloaded, and the model was trained from scratch. This approach eliminated the possibility of residual learning or cumulative biases influencing the results, ensuring the integrity of each experiment.

The findings from these experiments demonstrate that augmenting the dataset before splitting into folders is the optimal strategy for enhancing defect detection performance in PCB datasets. The normalized confusion matrix provided further substantiates this conclusion by visually emphasizing the improved accuracy and generalization achieved through this methodology.

3.9 Discussion

In this section, five different figures will be presented, depicting the precision, recall, mAP50, F1 score, and power score across multiple experiments. To facilitate optimal comparison, each graph will feature the best experiment without any augmentation applied to the dataset at the bottom, followed by the best experiments with augmentation displayed from the second from the bottom to the top. The remaining figures will highlight the best or one of the top-performing experiments from each batch. This arrangement allows for a clear visualization of the performance differences between the various augmentation strategies and their impact on model metrics.

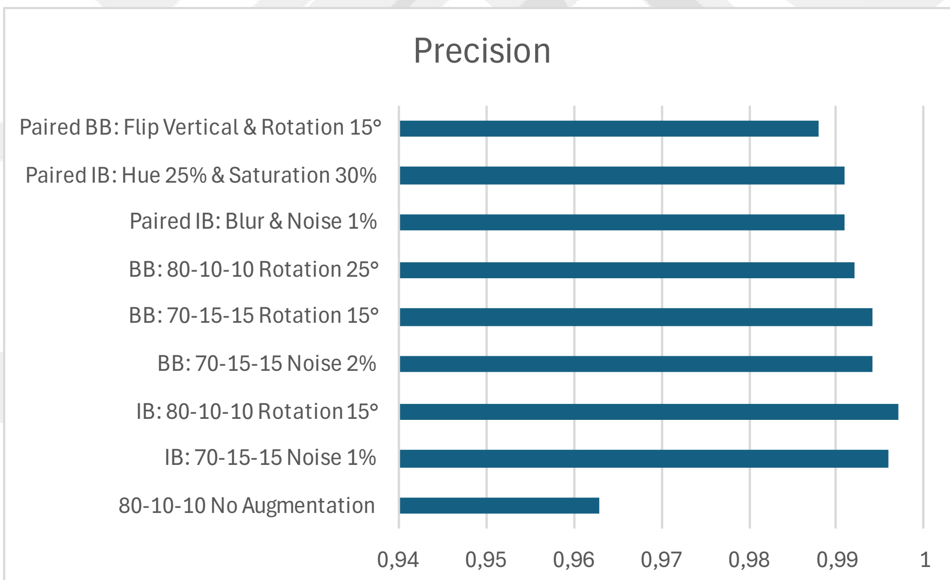

Figure 11 Precision Comparison

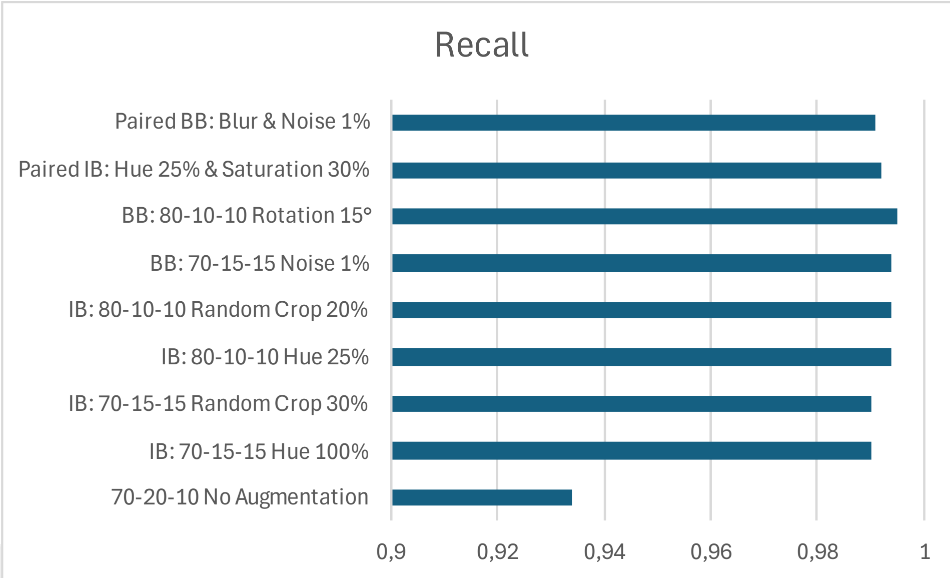

Figure 12 Recall Comparison

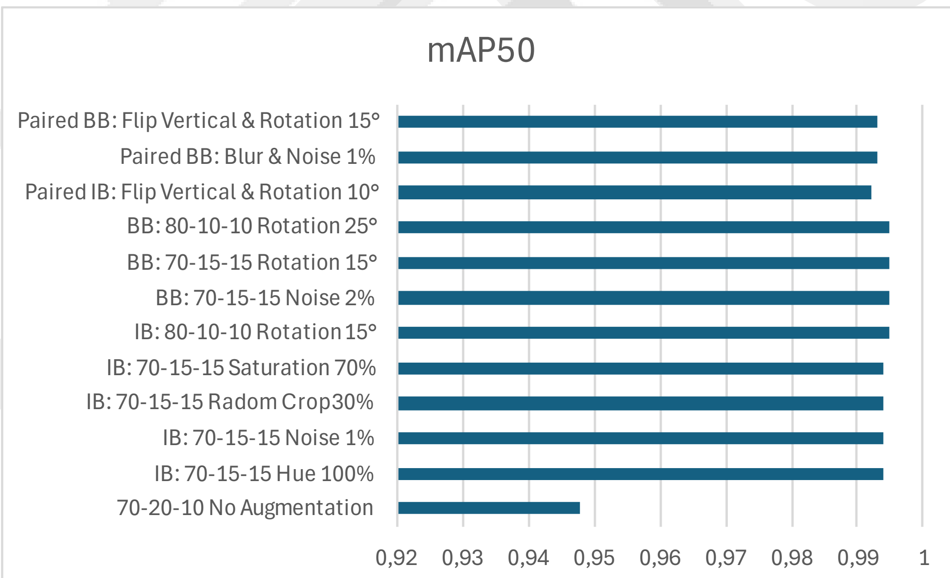

Figure 13 mAP50 Comparison

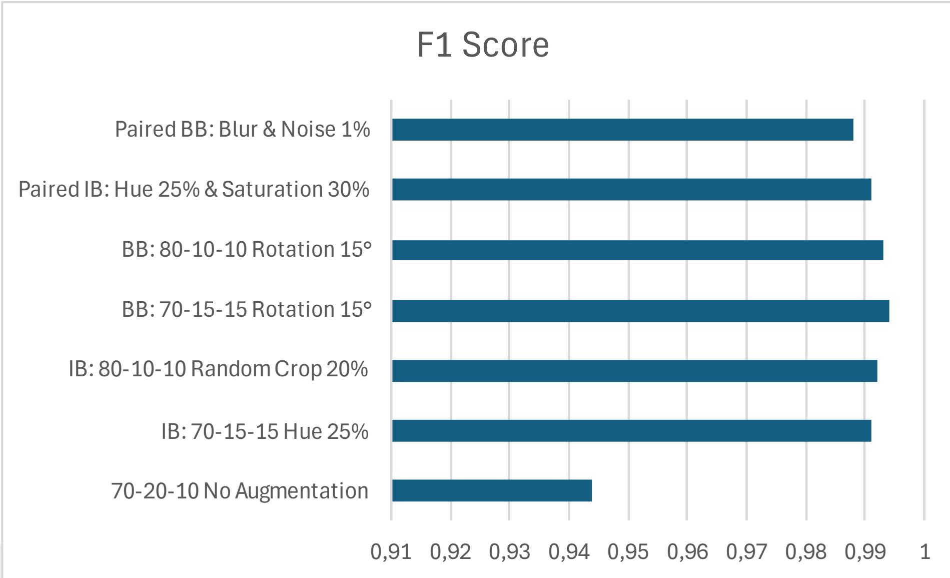

Figure 14 F1 Score Comparison

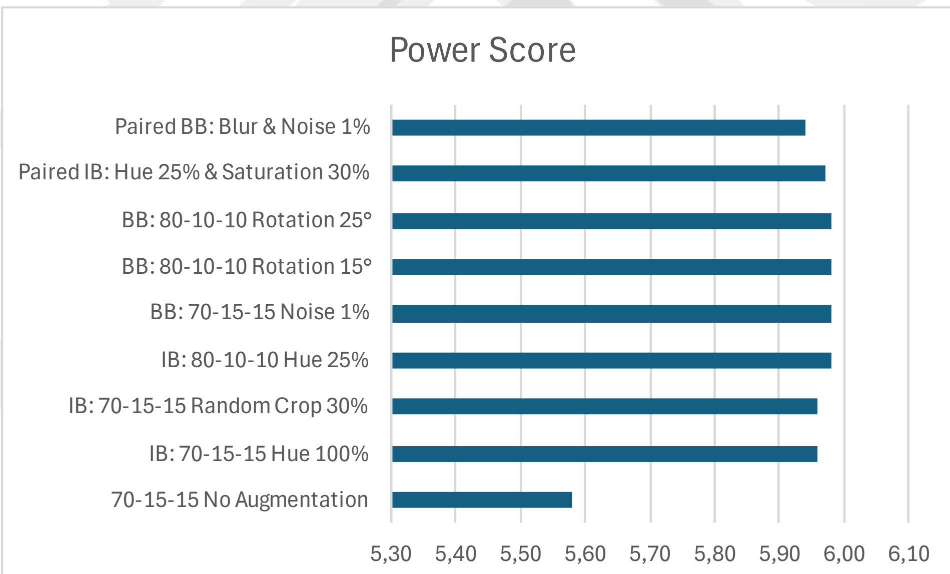

Figure 15 Power Score Comparison

The graphs distinctly illustrate a significant disparity between the experiment without augmentation and those incorporating augmentation, emphasizing that the choice of the most effective augmentation technique can substantially enhance performance outcomes. This reinforces the importance of strategic augmentation in optimizing model efficacy for tasks such as small object detection.

CHAPTER 4

CONCLUSION

This study demonstrates the significant impact of data augmentation techniques in enhancing defect detection for printed circuit boards (PCBs), particularly in scenarios with limited datasets. Among the techniques evaluated, moderate augmentations—such as small rotations, noise, and color adjustments—emerged as the most effective. These methods preserved the critical features of small components, enabling the YOLOv8 model to generalize efficiently and deliver high performance. In contrast, extreme augmentations, including large rotations, introduced distortions that negatively affected detection accuracy, reaffirming the importance of balance in augmentation strategies.

For the PCB manufacturing industry, the adoption of optimized augmentation approaches presents a promising avenue to enhance automated defect detection systems. These strategies can significantly reduce reliance on manual inspections, streamline quality control processes, and improve overall production efficiency. The findings also confirm and expand upon the insights from the literature review, emphasizing the crucial role of data augmentation in overcoming the limitations posed by constrained datasets. The alignment between experimental results and prior studies underscores the consistent value of carefully selected augmentations across diverse experimental setups.

By systematically validating the effectiveness of these techniques, this research not only corroborates existing knowledge but also provides a refined framework for leveraging augmentation to address the unique challenges of PCB defect detection. This work highlights that data augmentation is not merely a supplemental tool but a cornerstone for advancing AI-driven quality assurance in PCB manufacturing. Future research should explore the integration of augmentation techniques with other machine learning advancements, as well as their application to broader challenges in small object detection.

4.1 Answering the Research Questions

- **How do image-based augmentation techniques compare to bounding box-based techniques?**

Image-based augmentations exhibited superior performance by preserving contextual information and enhancing detection accuracy. In contrast, bounding box-based methods necessitated more precise tuning, with large rotations frequently resulting in decreased performance.

- **How does the amount of training data influence augmentation effectiveness?**

An increase in training data generally enhanced performance; however, the improvements diminished after a certain point. The performance improved when transitioning from the 70-15-15 to the 80-10-10 data split, but the improvement was not substantial. Therefore, the choice of split should depend on the dataset and specific requirements. If the larger training dataset in the 80-10-10 split significantly increases training time, opting for a smaller training data percentage, such as 70-15-15, may be more practical to achieve a balance between performance and efficiency.

- **Are certain augmentations more effective for detecting small components?**

Yes, moderate augmentations—such as small rotations, noise, and hue adjustments—were found to be the most effective for detecting small components, while extreme augmentations, like large rotations, adversely impacted detection accuracy.

- **What is the impact of color augmentation techniques?**

Color adjustments, including hue and saturation, had a positive effect on detection performance by enabling the model to focus on shapes and textures. This improvement in generalization was achieved without compromising accuracy.

Data Availability

- **Roboflow Link 1:**

https://universe.roboflow.com/pcbdddatasetsuralsezen/pcb_defect_detection_datasets_1_ural_sezen

- **Roboflow Link 2:**

https://universe.roboflow.com/pcbddatasetsuralsezen-fmwzm/pcb_defect_detection_datasets_2_ural_sezen

- **Google Drive SpreadSheet Link:**

https://docs.google.com/spreadsheets/d/1aSl7IyllrY-xtyWXTMntQXbMc6RZOWJs/edit?usp=sharing&ouid=104379159764244297260&rtpof=true&sd=true

REFERENCES

[1] D.I. Ural and A. Sezen, “Research on PCB defect detection using artificial intelligence: a systematic mapping study.” *Evolutionary Intelligence*, vol. 17, pp. 3101-3111, Apr. 2024.

[2] Y. Zhou, M. Yuan, J. Zhang, G. Ding and S. Qin, “Review of vision-based defect detection research and its perspectives for printed circuit board,” *Journal of Manufacturing Systems*, vol. 70, pp. 557-578, Oct. 2023.

[3] M. Borthakur, A. Latne and P. Kulkarni, “A Comparative Study Of Automated PCB Defect Detection Algorithms And To Propose An Optimal Approach To Improve The Technique,” *International Journal of Computer Applications*, vol. 114, pp. 27-33, Mar. 2015.

[4] M. Sohan, T. S. Ram and V. R. Reddy, “A Review on YOLOv8 and Its Advancements,” *Data Intelligence and Cognitive Informatics*, pp. 529-545, Jan. 2024.

[5] W. Huang and P. Wei, “A PCB Dataset for Defect Detection and Classification,” *Journal of Latex Class Files*, vol. 14, Aug. 2018.

[6] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, Jul. 2019.

[7] G. Lin, J. Jiang, J. Bai, Y. Su, Z. Su and H. Liu, “Frontiers and developments of data augmentation for image: From unlearnable to learnable,” *Information Fusion*, vol. 114, Feb. 2025.

[8] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao and F. Shen (2022, Apr.), “Image Data Augmentation for Deep Learning: A Survey,” *arXiv*, [Online], Available: https://arxiv.org/abs/2204.08610 [2024]

[9] K. Maharana, S. Mondal and B. Nemade, “A review: Data pre-processing and data augmentation techniques,” *Global Transitions Proceedings*, vol. 3, pp. 91-99, Jun. 2022.

[10] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, Dec. 2022.

[11] W. Li, “Detecting defects in PCB manufacturing: an exploration using Yolov8 deep learning,” *International Journal on Interactive Design and Manufacturing*, Jul. 2024.

[12] D. A. van Dyk and X. L. Meng, “The Art of Data Augmentation,” *Journal of Computational and Graphical Statistics*, vol. 10, pp. 1-50, Jan. 2021.

[13] J. Terven, D. M. C. Esparza and J. A. R. Gonzalez, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” *Machine Learning & Knowledge Extraction*, vol. 5, pp. 1680-1716, Nov. 2023.

[14] T. Diwan, g. Anirudh and J. V. Tembhurne, “Object detection using YOLO: challenges, architectural successors, datasets and applications,” *Multimedia Tools and Applications*, vol. 82, pp. 9243-9275, Aug. 2022.